MOON ECONOMICS INSTITUTE

WORKING PAPER 2022

Economic Values of Data and Data Flows, and Global Minimum Tax

Wendy C. Y. Li
Moon Economics Institute

MOON ECONOMICS INSTITUTE
https://moonecon.org/

Economic Values of Data and Data Flows, and Global Minimum Tax

Wendy C.Y. Li[1][2]

Moon Economics Institute

Revision Date: January 5, 2022

Abstract

In the digital era, data is the key input to a firm's production. Firms using data to organize production, notably Big Tech, enjoy higher productivity and market valuations. However, policymakers in many nations are concerned that Big Tech has not paid its fair share of taxes. A major reason for the gap between expectations and reality is that the value of data has not been capitalized into a firm's financial statements. I find that Big Tech possesses a tremendous value of data, and capitalizing their value of data can increase their profit rates significantly, which can easily meet the criterion of a 10% profit margin for the global minimum tax. For example, capitalizing Amazon's value of data can increase its average profitability by 17%, with an annual growth rate of 12.2%. For Big Tech as a whole, the average profitability during the same period of time increases 11.4%, with an annual growth rate of 2.8%. Moreover, the estimated global value of data is around three trillion dollars. Nonetheless, even if the global Internet traffic continues to grow, the global value of data may saturate, or possibly decline, if Big Tech continues to gain a higher share of global data at a rapid pace. This paper also presents the first estimated economic value of cross-region data flows, which is of the order of several hundred billion dollars. The distribution of this value, however, is very uneven due to the inhomogeneity of cross-region data flows. My analysis is useful for policymakers to understand how much economic value of cross-border data flows may be at stake, and to understand the related transaction costs that businesses may incur under a data governance and tax policy. The analysis is also important for firms to evaluate the impacts of global minimum tax and data localization policy.

JEL Codes: O3

[2] E-mail: wendyli@moonecon.org

1. Introduction

Big Tech and online platform companies, typically multinational firms, can operate mainly in one or a handful of countries while providing digital services worldwide. Under the current production-based tax system, these companies are known and often criticized for paying little or no tax in many countries where they serve online and from which they derive sales revenue from the local markets. To ensure those "digital firms" pay their fair shares in the countries where they earn revenue, 136 countries have recently agreed to impose a 15% global minimum tax on any multinational firm with annual revenue exceeding 20 billion euros and a profit margin exceeding 10% (Lawder and Thomas, 2021). However, as commented by Janet Yellen, the new agreement on the global minimum tax may not be applicable to one of the best-known Big Tech companies, Amazon, because its reported profit rate in 2020 was only 6.3% (France-Presse, 2021; World Report, 2021).

Big Tech has been reshaping how we live and how firms produce, and the pace has continued to accelerate during the Covid pandemic, as many more activities have moved online. The core of Big Tech companies relies on artificial intelligence (AI) algorithms and data. As AI algorithms become more affordable and adaptable, data becomes the key that determines the accuracy of the algorithms. In other words, data is the heart of Big Tech's competitiveness in the digital era. As a firm's key input, data are crucial for firms to innovate, produce, and compete in the digital era. Firms using data to organize production enjoy higher productivity and market valuation (Coyle and Li, 2021; Li and Chi, 2021). Due to the network effect of online platforms and the data network effect, Big Tech has been growing rapidly and can internalize the benefits of the externalities derived from data.

Big Tech and online platform companies have data-driven business models and heavily invest in data. For example, Google pays Apple traffic acquisition costs (TACs) for the rights to be the default search engine for the Safari web browser on iPhone, iPad, and Mac devices. TACs are the costs of accessing a pool of consumers and acquiring their data. While the growth of iPhone users has slowed significantly, the TACs that Google has paid to Apple have increased 12-fold within 5 years, and reached US $12 billion in 2019, which is equal to over 10% of Google's data-targeted advertising revenue in 2018 (Li et al., 2019).

Such data-driven business models allows them to serve the global market without facing traditional physical constraints, to easily scale up their businesses, and to enter adjacent industries and markets, as long as they can access the data needed. For example, Booking.com serves 137,791 destinations, with 28.9 million properties in 229 countries[3]. However, most of its operations are conducted in its Amsterdam headquarters. Another example is Oracle's recent US $28.3 billion purchase of Cerner, a large electronic health record vendor, to enter the healthcare service sector (Lohr, 2021). Electronic health records are regarded as a necessity for medicine to move into the digital age — a shift that, in the long run, should increase efficiency, reduce costs, and deliver better healthcare services. This indicates that data is the key for high-tech companies to enter a traditional industry and provide new services.

When Big Tech and online platforms concentrate their operations in one or a handful of countries, their daily operations inevitably involve tremendous cross-border data flows. This is one of the main reasons why many countries with net data export assert that Big Tech should pay

[3] Li, W.C.Y. et al. (2019). Value of Data: There's No Such Thing as a Free Lunch in the Digital Economy, VOX CEPR Policy Portal column article, July 23rd.

for the benefits derived from the data collected in their countries[4]. Data flows are important to digital trade; therefore, estimating the economic value of cross-border data flows is critical in providing useful information for global policymakers to devise and implement the new global minimum tax, a new tax system that will undoubtedly affect trade between countries.

Albeit a crucial asset to Big Tech and online platform companies, data has not been incorporated in a firm's income statement based on the current accounting standard. Instead, all investments in data are categorized as expenditures, a practice that can significantly underestimate the profit rates of digital firms. The lack of knowledge about the value of data associated with cross-border data flows is also an immediate obstacle to understanding data trade. As a result, it has been challenging for governments to use existing measures and address new issues introduced by firms that are data-intensive and have data-driven business models.

This situation is understandable, because the value of data and data flow was hard to measure until very recently. Through our series of studies on the value of data, we have found that the value of the tons of data collected from consumers and third-party sellers by Amazon's online retail markets is enormous (Li et al., 2019), and there is a law of Big Tech's value of data – when the global data flow increases five fold, Big Tech's value of data doubles (Li and Chi, 2021).

To our knowledge, no studies have measured the economic values of cross-border data flow, or the impacts of capitalizing the value of data on Big Tech's profitability and the existing incumbents in the sectors they entered. Therefore, this paper investigates how capitalizing the value of data of Big Tech affects their profits, and develops a methodology to measure the economic values of cross-border data flows around the world. Specifically, this paper seeks to

[4] Coyle, D. & Li, W.C.Y. (2021). The Data Economy: Market Size and Global Trade, ESCoE Discussion Paper, September.

answer the following questions: What is the economic value of global data flow? How does the capitalization of data affect Big Tech's profitability? What are the economic values of cross-border data flows? What is the distribution of the economic values of cross-border data flows?

Our data sources include Big Tech's10K reports. The data cover the period between 2002 and 2017. In addition, the data source for global data flow is Cisco System, which covers the period from 2002 to 2017. Finally, the data sources for cross-border data flows are the International Telecommunication Union (ITU) and TeleGeography, covering 2015 to 2020.

Our key findings are shown below. First, Big Tech possesses a tremendous value of data, and data can increase Big Tech's profitability significantly. For example, capitalizing Amazon's value of data can increase its average profitability by 17%, with an annual growth rate of 12.2%. For Big Tech as a whole, the average profitability during the same period of time increases 11.4%, with an annual growth rate of 2.8%. Second, the entry of a data-driven online platform can negatively impact the profitability of traditional firms in the same sector. Third, the global value of data is estimated to be close to three trillion US dollars. However, even if the Internet traffic continues to grow, the global value of data may saturate, or possibly decline, if Big Tech continues to gain a higher proportion of their shares at a rapid pace. Fourth, the combined value of cross-region data flows is estimated to be of the order of several hundred billion US dollars, but the distribution is very uneven across the globe, due to the inhomogeneity of the cross-region data flows. The cross-region data flow between North America and Europe has the highest value. This study is the first that has developed a methodology for estimating the economic values of data associated with cross-border data flows and has demonstrated the empirical results.

The rest of paper proceeds as follows. Section 2 introduces the empirical methodology. Section 3 describes the data sets used. Section 4 shows the empirical analysis results. Section 5 concludes.

2. Empirical Methodology – Economic Values of Data and Cross-border Data Flows

Measuring the value of data is hard both conceptually and empirically. Recent studies have enhanced our understanding of the concept of the value of data (Li et al., 2019; Coyle et al., 2020; Li and Chi, 2021; Coyle and Li, 2021). As in the common issue with measuring intangibles, data is mainly produced for firms' own use, and there is no arm-length market available for most of them. This problem is especially severe because firms do not know how to value their data, and because governments and consumers have increasingly been concerning privacy and data security. There are several approaches to measure the value of data (Li et al., 2019; Coyle et al., 2020). However, most approaches do not define what the value of data is, nor do they provide a methodology for estimating the depreciation rate of the value of data either. For example, the labor cost approach proposed by Statistics Canada uses the salary of data scientists and workers (such as data labeling workers) as the value of data, and the method arbitrarily assigns the service life of data (Statistics Canada, 2019). In addition, most approaches cannot measure the impact of incorporating the value of data on a firm's profitability because they cannot estimate the firm-specific depreciation rate of the value of data, which is important to the problem because it affects the estimation of a firm's profits. Moreover, most approaches cannot demonstrate any connection between the explosive growth of global data flow to an economic value, a measurement conundrum that has been recognized by all recent studies, such as Coyle et al. (2020) and Tomiura et al. (2020).

To measure the value of data, we adopt the Li and Chi (2021) approach to estimate the value of data. This approach defines what the value of data is, provides a methodology to estimate the depreciation rate of the value of data, and so far is the only methodology that provides the solution to the measurement conundrum for linking the explosive global data flow to an economic value. Firms can use data to derive firm-specific knowledge, which can be measured by their organizational capital, the accumulated information or know-how of the firm (Li and Chi, 2021; Coyle and Li, 2021; Prescott & Visscher, 1980). As shown in the classical information pyramid, data is a key part of organizational capital, the firm-specific knowledge, and the value of data lies in the firm-specific knowledge derived from the use of data (Li and Chi, 2021; Coyle and Li, 2021). In other words, the value of data is a firm's specific knowledge derived from data (e.g., transaction data) which guides a firm to produce, compete, and grow. Firms with an advantage in data can better produce, compete, and grow.

Li and Chi (2021) discover that, when the global data volume increases by five folds, the Big Tech's value of data doubles. This is a clear empirical relationship that connects the global data flow to Big Tech's value of data (Figure 1 (c)). The Big Tech companies included in their study are Microsoft, Amazon, Apple, Google, Facebook, Alibaba, and Tencent, which are also among the global top-ten companies in 2019.[5] This is, so far, the only approach that solves the measurement conundrum of linking the exponential growth of global data flow to an economic value, and indicates that, not only has Big Tech created a tremendous amount of value from global data, but they are also the ones benefiting from the great economic opportunities generated by the

[5] Note that the Chinese government has adopted a tougher approach towards Chinese Big Tech, including Alibaba and Tencent, since 2020.

explosive global data growth. In addition, the clear empirical relationship between Big Tech's value of data and the global data flow can provide a useful tool to estimate the economic value of cross-border data flow.

1) Table 1: Global Top Ten Most Valuable Companies

Date: June 18, 2019 (for market cap)

Ranking	Company	Businesses	Market Cap†	Organizational Capital*[6] /Market Cap
1	**Microsoft**	Internet/Online Platform	1,036	15.38%
2	**Amazon**	Online Platform	936	12.19%
3	**Apple**	Internet/Online Platform	913	2.08%
4	**Google**	Online Platform	767	4.33%
5	**Facebook**	Online Platform	538	2.00%
6	Berkshire Hathaway	Financial	505	
7	**Alibaba**	Online Platform	431	1.32%
8	**Tencent**	Online Platform	403	1.22%
9	Johnson & Johnson	Pharmaceutical	372	
10	Visa	Financial	370	

† Unit: US $1 billion
* Source: Li and Chi (2021)

[6] This column indicates that firms with a higher degree of organizational capital intensity are also more valuable. This is consistent with the main finding in Eisfeldt and Papanikolaou (2013).

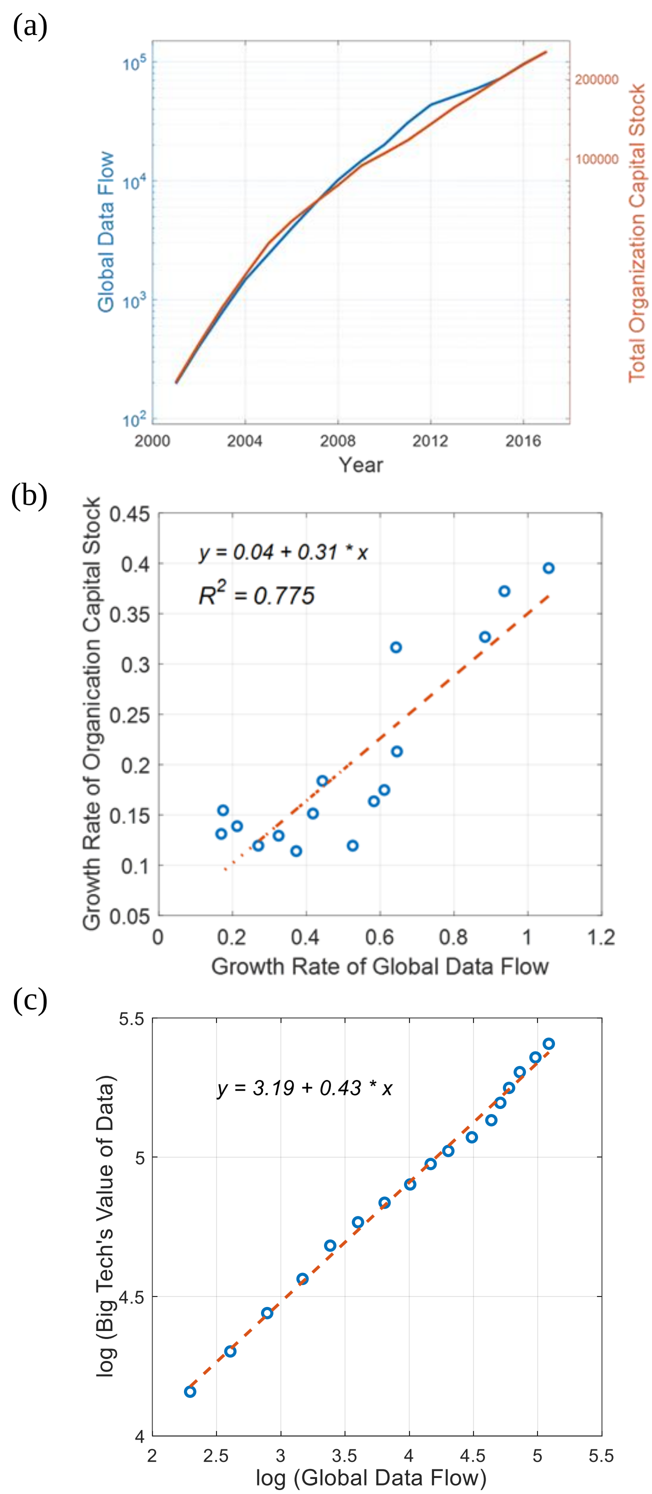

Figure 1. Big Tech's Combined Organizational capital vs. Global Data Flow.

Source: Li and Chi (2021)

Companies include Microsoft, Amazon, Apple, Google, Facebook, Alibaba, and Tencent.

2.1 Economic Values associated with Global Data Flow and Cross-border Data Flow

2.1.1 Law of Value of Data

Li's law of value of data can predict the value of Big Tech's data from the volume of global Internet traffic (Li and Chi, 2021). The law states that when the global data flow increases by five folds, Big Tech's value of data doubles, and the exact expression can be written as

$$\log K = 3.19 + 0.43 \times \log V$$

where V is the global Internet traffic (in PB/month) and K the combined value of data (in million US dollars) possessed by Big Tech, including Alphabet, Amazon, Microsoft, Apple, Meta (formerly known as Facebook), Alibaba, and Tencent (see Li and Chi, 2021).

Because the Internet traffic data published by Cisco are available only through 2017, and because the Internet traffic has been following a persistent growth path in the past two decades, we can extrapolate the time series data to the following several years (see **Figure 2**) to obtain a rough estimate of the Internet traffic and calculate the approximate combined value of data for the seven big tech firms.

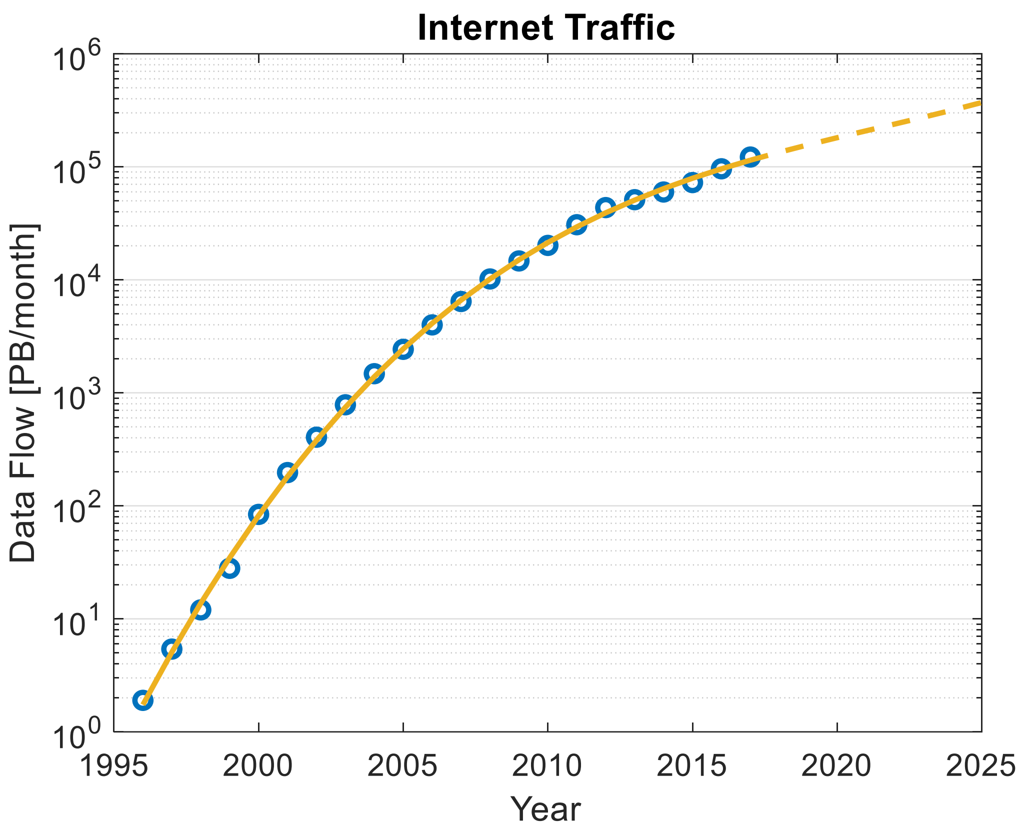

Figure 2. Internet traffic data, as reported by Cisco.

2.1.2 Global Value of Data

To understand the magnitude of global value of data, we first estimate the ratio between the value of data in the world and the value of data occupied by those of the top seven big tech firms. While it is conceptually feasible to calculate this ratio if one has the organization capital investment data for every company, it is difficult in practice because only public firms publish SG&A data. The information about the investment in organizational capital for private firms requires a large-scale survey that takes a significant amount of efforts by all the countries involved.

Due to the above reason, we use the market capitalization as a measure to estimate "the-world-to-big-7" ratio, assuming that organization capital of a firm is roughly proportional to its market capitalization. The calculation of the ratio only involves the mean value for the world and that for the big 7, and variability from this assumption can be reduced by averaging. Based on the publicly available data, we find that the combined market cap owned by the 7 Big Tech companies gradually increases its share of the global market cap. **Figure 3** depicts this trend and a second-order polynomial fit. It is clear that, since 2014, these seven companies have made tremendous ground in market shares. If this trend continues, the world's market cap will only be 6.4 times of the combined market cap of the 7 big tech firms in 2024. With this ratio, the global value of data can then be estimated by multiplying the value of data for the top seven big tech firms by this ratio.

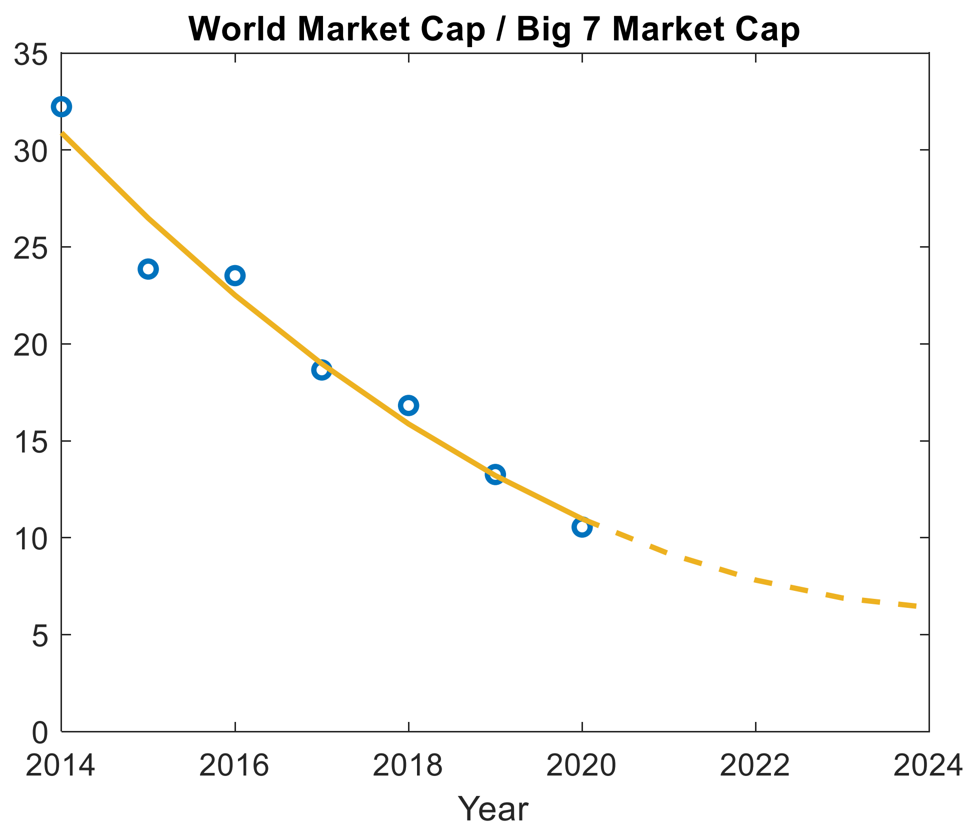

Figure 3. Ratio between the market cap for the world and that for the top seven big tech firms.

2.1.3 Economic Values of Cross-Region Data Flows

As we have demonstrated in the connection between Big Tech's value of data and the global Internet traffic, a topic of interest is the distribution of the value of data around the world. In particular, data from the International Telecommunication Union (ITU) and TeleGeography show that cross-border data flows have increased substantially from 2015 to 2020 (UNCTAD, 2021), and how much the value of data is associated with these cross-border data flows has become an important question for global data governance, trade, and tax policies.

If we consider that the value of data is correlated with the quantity of data, as suggested by the law of the value of data, we can estimate the value of data associated with each cross-region data flow based on the magnitude of the data flow and the global value of data. In this study, the analysis is based on the published data for cross-region data flows without the information of the flow direction. The same approach can be applied to understanding the flow of the value of data and who benefits from the data flow once the microdata for the direction and ownership of Internet traffic become available.

2.2 Profitability

Hulten and Hao (2008) and Hulten (2010) calculated the impacts of capitalizing intangibles on a firm's profit and earning per share. When they calculated the profits of firms, they assumed the lag of intangible investments to be zero and used the current-period investment in the intangible capital plus the return to the intangible capital as the current revenue. Additionally, they assumed an ad hoc depreciation rate for different types of intangible capitals.

As discussed in Li and Chi (2021), using an ad hoc-fixed depreciation rate for an intangible capital across the board has three major problems. First, it cannot reflect the impact of new business innovations on the organizational capital of existing firms, a problem that is especially serious in the digital era. Second, we need a firm-specific depreciation rate of an intangible capital to estimate a firm's profit. Finally, there is a difference between national accounting (Hulten and Hao, 2008; Hulten, 2010) and firm-level accounting in capitalizing intangibles. To estimate an individual private firm's profit, we do not add the combined value of the current-period investment in intangibles with the return to intangibles to the current-period revenue. Like R&D assets, data as an asset has two benefits to a firm: sales increase and profit increase – both should last for more than a year. The current sales of the firms already reflect the impacts of the previous and/or current investment results embodied in the products and market competitiveness. This is similar to how an accountant calculates a firm's profit when there is an investment in a new machine: in this case, the accounting department will not add the investment in machine plus its return to existing revenue when calculating the profit.

3. Data

Following earlier related studies, we use the selling, general, and administrative (SG&A) expense as a proxy for a firm's investment in organizational capital (Lev and Radhakrishnan, 2005; Eisfeldt and Papanikolaou, 2013; Li, 2015, 2016a; Li et al., 2019). Firms report this expense in their annual income statements, and it includes most of the expenditures that generate organizational capital, such as employee training costs, brand enhancement activities, consulting fees, and the installation and management costs of supply chains. Specifically, we examine the data for Microsoft, Amazon, Apple, Alphabet, Meta, Alibaba, and Tencent. Our SG&A data are from firms' public income statements, and cover the years 2002 to 2017. In addition, Cisco System provides the data source for the global Internet traffic from 2002 to 2017. The data for cross-border data flows between 2015 and 2020 come from the International Telecommunication Union (ITU) and TeleGeography.

4. Empirical Results

4.1 Economic Values of Cross-border Data Flows

Based on the methodology described in Section 2.1.2, we find that the global value of data is at the order of several trillion dollars. For example, the estimated global value of data is $2.98 trillion in 2020. Section 4.3 describes the estimated values during 2014-2017 and the projected values in later years in detail. Because the value of data is correlated with Internet traffic, as indicated by the law of the value of data, the value associated with each subset of Internet traffic, such as cross-region data flows, can be estimated by the proportion of the subset.

The data flows as shown by UNCTAD (2021) with data sources from International Telecommunication Union (ITU) and TeleGeography concern about the international Internet traffic only. To have the first-order estimate of the combined value of data associated with the cross-border data flows, we use the ratio between the international trade and the world GDP, which is 18% in recent years. This value of data is further divided by the various cross-region data flows according to the data volume.

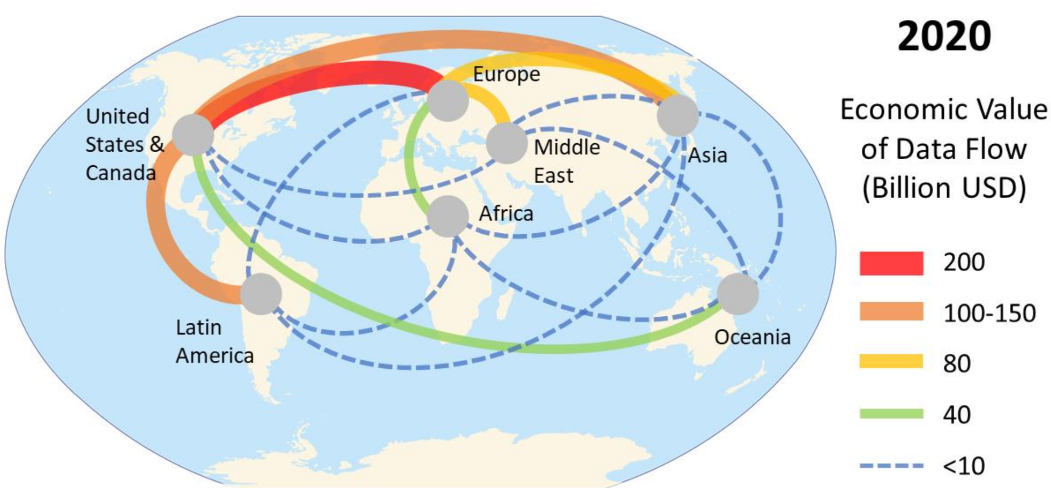

Figure 4. Economic value of cross-region data flow.

Figure 4 shows the estimates of the economic value associated with each cross-border data flow during 2020, based on the cross-region data flows indicated in the report by UNCTAD (2021) and our estimates of the economic value of data. Here the data flow only refers to the amount of data per month without indicating in which direction they flow.

To understand how this economic value of cross-border data flow, data trade, is compared with the size of traditional trade goods and services in 2019, the estimated trade goods and services is US $1.1 trillion, according the Office of the U.S. Trade Representative (USTR, 2020). Therefore, when setting the data governance policy, policymakers need to recognize the magnitude of the economic value of cross-border data flow related to their regions. In particular, when imposing the data localization policy, policymakers need to understand how much economic values of cross-border data flows will be interrupted, as well as the related extra transaction costs will occur for businesses. From the business side, this piece of information will be important when they evaluate the impacts of global minimum tax and data localization policy. Because there is no data on the ownership and utilization of those cross-border data flows at the firm level, we cannot calculate the size of an individual firm's data trade or its gains from data trade.

It is noted here that the estimated economic values of cross-border data flows for non-Big Tech firms may be overestimated if we apply higher depreciation rates. Li (2015, 2016) and Li and Chi (2021) find that in China, Germany, S. Korea, Japan, and the U.S., industry leaders and sector leaders normally have lower depreciation rates of intangible capital, including the value of data.

4.2 Profitability

Having a complete picture of how firms benefit from data requires us to calculate the value of data owned by firms and understand how incorporating the value of data into a firm's financial statements affects its profitability. For the top seven big tech firms, I examine the annual OC investment (SG&A-R&D, denoted as A_t where t is the index for year) and the amount of depreciated OC per year (denoted as B_t). The depreciation rate of organizational capital, δ, for each of the seven firms is calculated using the Li and Hall (2020) method and is included in **Table 1**.

Table 1. Organization capital depreciation and profit for the top seven big tech firms for data during years 2010-2017

	Delta (OC)	$(A_t - B_t)/Sales_t$ mean value	$A_t - B_t$ mean value*	$\Sigma(A_t - B_t)$ mean value*	$\Sigma(A_t - B_t)/\Sigma(Sales_t)$
Alibaba	47.7%	0.174	1155	3447	0.154
Tencent	46.1%	0.103	1288	4677	0.110
Microsoft	12.9%	0.085	6662	10155	0.137
Apple	63.2%	0.025	3865	21487	0.026
Facebook	52.2%	0.152	1966	5841	0.141
Amazon	43.6%	0.167	16504	59878	0.132
Alphabet	42.4%	0.095	5927	31292	0.100
Average		0.114	5338	32597	0.114

*In million USD.

The quantity $(A_t - B_t)$ is the profit resulting from capitalizing organizational capital (OC). **Figure 5** shows the time series of this profit as well as its accumulated value. The OC profit has an increasing trend for almost all top firms, except for Microsoft where the OC profit per year kept steady but remained in the lead throughout this time interval. The figure also indicates that it is common for these top firms to double the accumulated OC profit in two years.

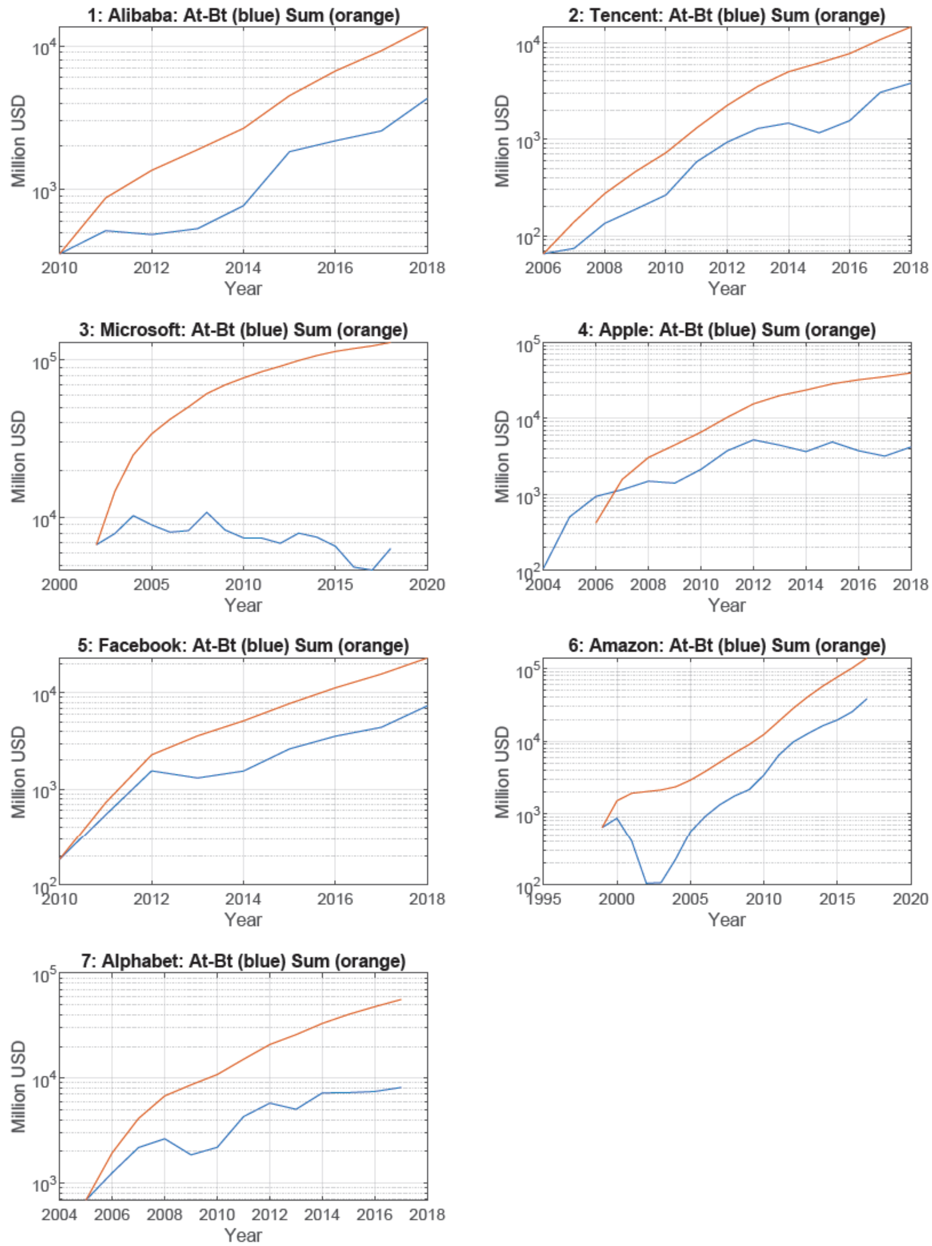

Figure 5. Time series of OC profit (blue) as well as its accumulated value (orange).

The profit rate increase due to capitalizing OC can be calculated by the ratio of $(A_t - B_t)$ to sales. **Figure 6** shows the annual OC profit rate as well as this rate based on accumulated sales and OC profits. The OC profit rate for most of these top firms was approximately 10% or slightly higher, since 2013. The two exceptions are Apple, which maintained the OC profit rate below 5%, and Amazon, whose OC profit rate was rising toward 20% in 2017 for the data set examined.

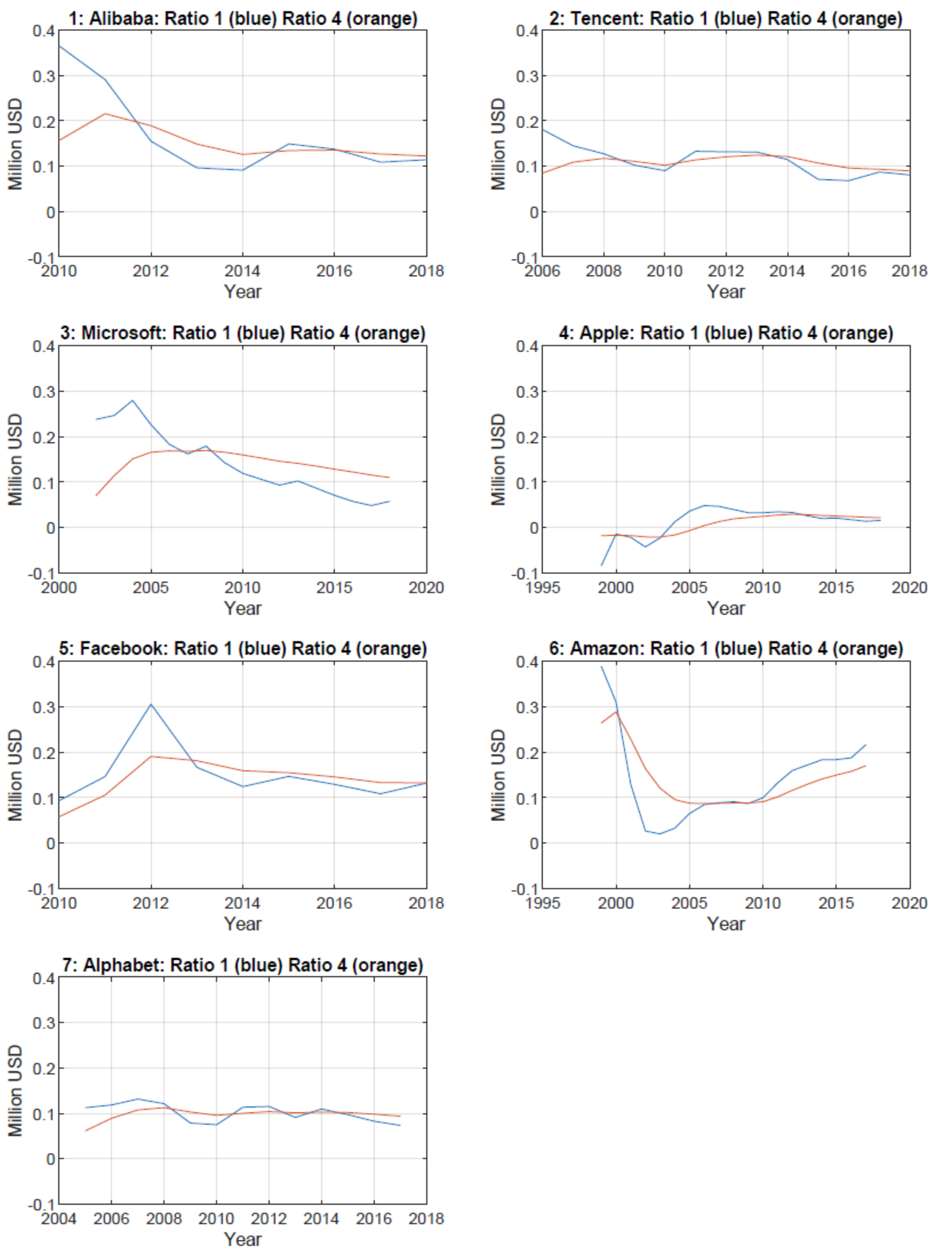

Figure 6. OC profit rate: annual (blue) and average (orange).

Table 2 lists the average values of the growth rates of OC profit and profit rate for the top seven big tech firms. On average, the annual OC profit for these top firms grew 33% during 2011-2017, including growth rates of over 40% for Amazon and Tencent and an unusually fast growth rate of 73% for Facebook.

Table 2. Growth rates of OC profit and profit rate.

Growth Rates (during 2011-2017)

	Delta (OC)	Growth rate of $(A_t - B_t)/Sales_t$	Average Annual Growth rate of $A_t - B_t$
Alibaba	47.7%	-0.108	0.381
Tencent	46.1%	0.028	0.489
Microsoft	12.9%	-0.116	-0.058
Apple	63.2%	-0.107	0.111
Facebook	52.2%	0.121	0.733
Amazon	43.6%	0.122	0.429
Alphabet	42.4%	0.022	0.248
Average		-0.005	0.333

It is worth noting that, because the depreciation of organizational capital directly reduces the OC profit, knowing the OC depreciation rate (δ) for the firm of consideration is essential. This point can be well demonstrated by the case study of Airbnb's impact on Marriott's organization capital. As shown by Li and Chi (2021), the OC depreciation rates of Marriott before and after the entry of Airbnb to the hospitality industry are 45% and 55%, respectively. The increase in Marriott's δ was a consequence of the "creative disruption" by Airbnb in the same industry sector that made Marriott's business model relatively outdated.

Figure 7 shows Marriott's OC profit rate estimated by the same approach as above. It is clear that, right after Airbnb entered the industry in 2008, the rise in Marriott's δ drove the OC profit rate down to the negative domain. Marriott's OC profit did not recover until 2015, when it acquired Starwood and became the largest hotel group in the world.

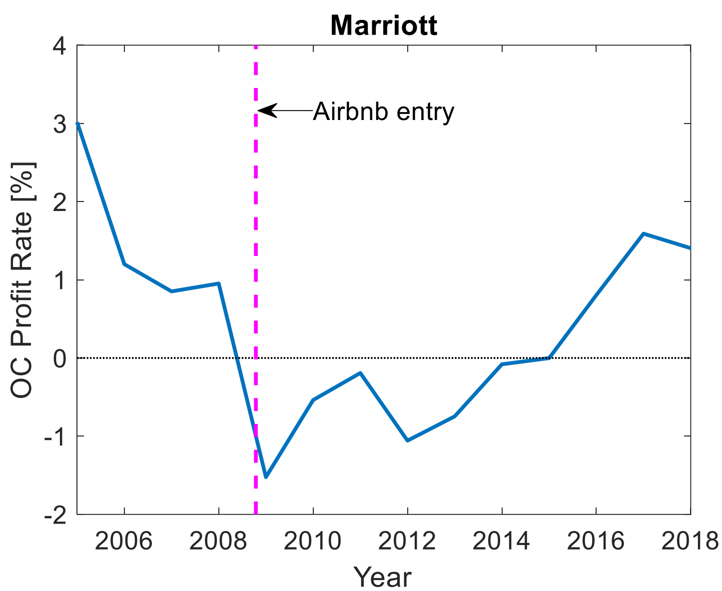

Figure 7. Variations in Marriott's OC profit rate before and after the entry of Airbnb.

4.3 Global Value of Data

Figure 8 shows the estimated global value of data from 2014 to 2017. Different from expectations based on the growing Internet traffic, these estimated values were already on a slightly declining trend, due to the fact that Big Tech companies were rapidly gaining shares of worldwide market cap during this period of time. We consider the following different scenarios for projecting the global value of data in the years after 2017:

i. If both the Internet traffic and the combined OC stock of Big Tech followed the trends calculated based on the data before 2018, the projected global value of data continues to decline (blue curve). The global value of data in 2024 will be reduced to approximately half of the value in 2014.

ii. If we consider a different scenario in which all conditions are identical to those in (i), but the growth rate of Internet traffic is doubled starting from the COVID pandemic in 2020, the global value of data will stop falling in 2022, and rise again (green curve). This recovery would be due to the rapid growth in Internet traffic.

iii. A third scenario considers that all conditions are identical to those in (i), except that the combined market cap of the 7 top Big Tech firms will be kept at 10% of the world value (see the market cap ratios for earlier years). In this case, the global value of data will start rising after 2021 at a faster pace than the recovery in scenario (ii) (red curve).

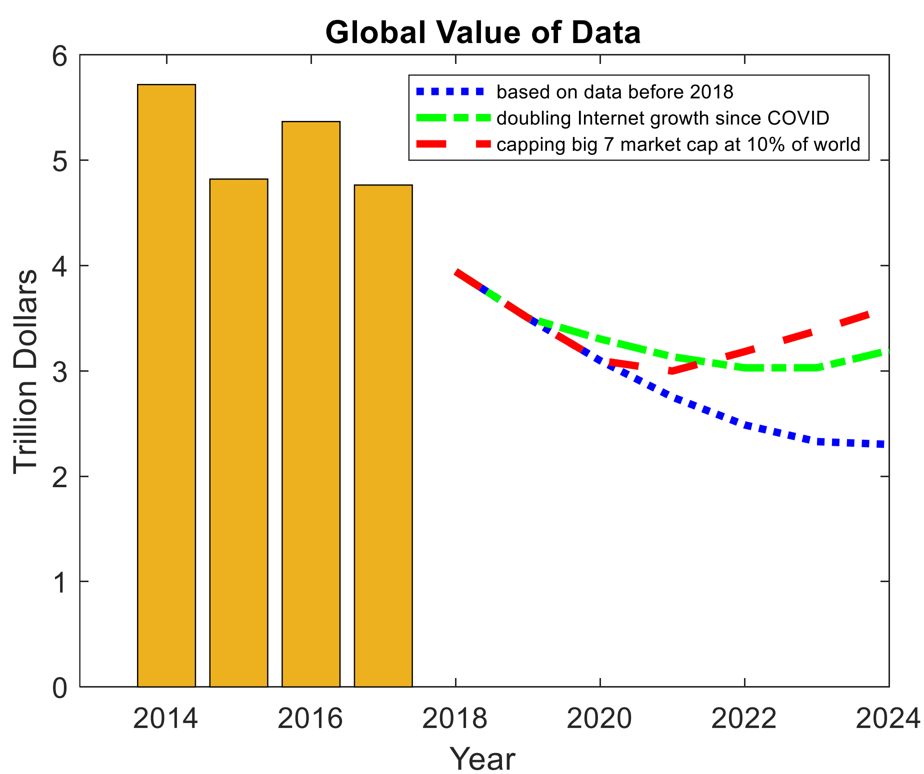

Figure 8. Estimated global value of data during 2014-2017, and projected values in three different scenarios.

This reverse trend suggests an interesting phenomenon in which, unlike the law of Big Tech's value of data, the growth in the global value of data can be affected by the degree of diversity in accessing the data and does not necessarily follow the growth in Internet traffic. In addition, the global data flow after 2017 is estimated based on the previous data from Cisco. According to UNCTAD, the global data flow has been undergoing explosive growth after 2018, particularly after the start of Covid-19 in 2020, and the data growth has further increased due to the accelerated digitization of the world's economy (UNCTAD, 2021). The new development might delay the saturation of the global value of data as predicted above, but the analysis still suggests that the growth in the value of data created can be limited and even reduced when the data access and utilization is mostly concentrated and controlled by Big Tech and is limited for other firms. Therefore, it is of public interest if supporting institutions and rewarding mechanisms can be established for fostering an ecosystem for data sharing.

5. Conclusion

Big Tech has been reshaping not only how we live but also how firms produce, and this change has been accelerating during the Covid pandemic when even more activities are being moved online. The core of the "digital" companies relies on algorithms and data. When AI algorithms become more affordable and adaptable, data become the key determining the accuracy of the algorithms. Thus, data is the heart of Big Tech's competitiveness in the digital era. Aided by the network effect of online platforms and the data network effect, Big Tech has been growing rapidly and can internalize the benefits of the externalities derived from data. In addition, the data-driven business model allows Big Tech firms to serve the global market by concentrating their operations in one or a handful of regions while scaling up their businesses easily, as long as they can access the local data.

On the other hand, Big Tech's data-driven operation model allows these firms to pay zero or little tax under the current production-based tax system. One well-publicized criticism to Big Tech in many countries is its collection of data from people and businesses to help the firm earn tremendous revenues from local markets. Nonetheless, it has been difficult for governments to curtail the negative impacts generated by firms with data-intensive and data-driven business models. A key problem is that governments do not know the magnitude of the economic values of data and cross-border data flows, which can significantly affect a firm's profitability and data trade. As demonstrated by the recent international agreement on the global minimum tax with a threshold of 10% profit margin, companies like Amazon can become slippery fish when the value of data is not considered in the firm's profitability.

This paper presents a methodology that can examine the impacts of capitalizing the value of data on a firm's profitability. We find that capitalizing Amazon's value of data can increase its average profitability by 17%, with an annual growth rate of 12.2%. For Big Tech as a whole, the average profitability during the same period of time increases 11.4%, with an annual growth rate of 2.8%. This means that Big Tech's profitability can easily meet the criterion of 10% profit margin when its value of data is capitalized.

This paper also presents a methodology to estimate the economic value of cross-border data flow around the world. To our knowledge, this is the first empirical estimate of the data trade between continents, and it shows that the economic value of cross-border data flow is enormous. Our calculations indicate that these cross-region data flows can translate to hundreds of billions of U.S. dollars of value of data per year. This economic value of cross-border data flow based on the combined values for two-way internet traffic effectively is data trade. It is worth noting that the magnitude of this data trade today is nontrivial when comparing it with the size of traditional trade goods and services (for example, the traditional trade goods and services between the US and Europe were worth $1.1 trillion in 2019, according to the Office of the U.S. Trade Representative). Therefore, it is helpful for policymakers to consider the economic value of cross-border data flows to and from their regions when formulating the data governance policy. In particular, when imposing mandates for data localization, one should not lose sight of the economic value of the cross-border data flows that may be interrupted, and the ensuing extra transaction costs that may incurred for businesses. From the business perspective, understanding the economic value of cross-border data flows can aid in the evaluation of the impacts of the global minimum tax and data localization policy.

Last but not least, we find that, unlike the law of Big Tech's value of data, the growth in the *global* value of data does not necessarily follow the growth in the Internet traffic and can be affected by the degree of diversification in data use. During 2014-2017, the estimated global value of data has already shown a slight declining trend due to Big Tech's rapid gain in shares. Although this trend might have changed in 2020-2021, during which the COVID-19 pandemic forced a dramatic increase in online activities, the implications from the analysis based on pre-pandemic data still hold. When the data access and utilization is heavily concentrated and controlled by Big Tech and is limited for other firms, the growth in the global value of data created can be suppressed and even become negative.

Therefore, a more visionary topic for policymakers is to explore possible institutions and rewarding mechanisms for incubating a data-sharing ecosystem in order to further unlock the power of data. As Hal Varian (Chief Economist of Google) rightly pointed out at the 2021 ASSA

Meeting, data is a club goods, which is excludable but non-rival. The Covid-19 pandemic has pushed all firms to accelerate their digital transformation, which relies on data. Hence, more firms have started to invest in data collection (Yi, 2021). However, because of the data network effect and the network effect of online platforms, it is hard for non-Big Tech firms to overcome their disadvantage in data by collecting data alone. The current discussions in global data governance policies, such as data localization, are focused mainly on privacy and data security. These are important issues, but it is also necessary to avoid the pitfalls if the actual measures for these policies in fact reinforce the data advantage of Big Tech that controls the gateway of data and has more resources to deal with the compliance costs, consequently worsening the data inequality among firms and between countries.

References

Coyle, D., Diepeveen, S., Wdowin, J., Kay, L., & Tennison , J. 2020. The Value of Data: Policy Implications, February, https://www.bennettinstitute.cam.ac.uk/media/uploads/files/Value_of_data_Policy_Implications_Report_26_Feb_ok4noWn.pdf

Coyle, C. & Li, W.C.Y. 2021. The Data Economy: Market Size and Global Trade, ESCoE Discussion Paper, September.

Eisfeldt A., & Papanikolaou, D. 2013. Organizational capital and the Cross-Section of Expected Returns, *Journal of Finance*, 4, August, pp: 1365-1406.

France-Presse, A. 2021. Yellen 'Not Certain' Amazon Would Pay Under Global Tax Agreement, https://www.courthousenews.com/yellen-not-certain-amazon-would-pay-under-global-tax-agreement/

Hulten, C. & Hao, X. 2008. What Is A Company Really Worth? Intangible Capital And The "Market To Book Value" Puzzle, NBER Working Paper #14548, December.

Hulten, C. 2010. Decoding Microsoft: Intangible Capital As A Source of Company Growth, NBER Working Paper #15799, March.

Lawder, D. & Thomas, L. 2021. U.S. Drops Tariff Threat in Digital Tax Transition Deal with European Countries, Reuters.com, October 21.

Lev, B., & Radhakrishnan, S. 2005. The valuation of organization capital. In *Measuring Capital in a New Economy*, Corrado, C., Haltiwanger, J., Sichel, D. (eds). Chicago: National Bureau of Economic Research and University of Chicago Press: Chicago, IL; 73–99.

Li, W.C.Y. 2015. Organizational Capital, R&D Assets, and Offshore Outsourcing, 2015 ASSA/SGE Sessions Proceeding Paper.

Li, W.C.Y. 2016a. Offshore Outsourcing and U.S. Innovation Capacity, Working Paper, Presented in the 2016 National Bureau of Economic Research Summer Institute/CRIW Workshop, July.

Li, W.C.Y. 2016b. New Technology Indicator for Technological Progress, Working Paper, Presented at the 34th IARIW General Conference at Dresden, Germany.

Li, W.C.Y. 2020. Depreciation of Business R&D Capital, Review of Income and Wealth, https://doi.org/10.1111/roiw.12380.

Li, W.C.Y. 2019a. Value of Data: There's No Such Thing as a Free Lunch in the Digital Economy, VOX CEPR Policy Portal column article, July 23rd.

LI, W.C.Y. 2019b. Value of Data: There's No Such Thing as a Free Lunch in the Digital Economy, RIETI discussion paper 19-E-022.

Li, W.C.Y. & Chi, P.J. 2021. Online Platforms' Creative "Disruption" in Organizational Capital – The Accumulated Information of the Firm, Moon Economics Institute Discussion Paper, No. 1, August.

Lohr, S. 2021. Oracle takes a big move toward health with a deal to buy Cerner for $28.3 billion, *the New York Times*, December 20. https://www.nytimes.com/2021/12/20/technology/oracle-cerner-health-records.html

Prescott, E., & Visscher, M. 1980. Organization Capital, *Journal of Political Economy*, 88, pp: 446-461.

Tomiura, E., Ito, B., & Kang, B. 2020. Cross-border Data Transfers under New Regulations: Findings from A Survey of Japanese Firms, *VOX CEPR Policy Portal* column article, March 14.

United Nations Conference on Trade and Development. 2021. Digital Economy Report 2021 –

Cross-border Data Flows and Development : From Whom The Data Flow, September.

Yi, C.Y. 2021. Firms Change Strategies, Build Their Own Customer Database. *The World Journal*. December 20.